Dear Editor,

We thank you for the opportunity to submit a revised version of our manuscript. In this version, we have addressed all comments from the reviewers and added new layers of evidence to support our main findings. Specifically, we have included data from a new experiment where we quantify how the populations of *Tetranychus urticae* and *T. evansi* modify the expression of defense genes on plant hosts, which is our proxy for niche modification. Results from this experiment support our previous suggestion that niche preemption was the main mechanism underlying the changes in coexistence outcomes between the two spider mite species. Due to the addition of this experiment, we have also added a new co-author, Mariya Kozak, who performed the lab work associated with this data.

Moreover, in response to the constructive comments provided by the two Reviewers and Associate editor, we emphasize two important points that were improved in this revision. First, following a comment by Reviewer 2, we re-analyzed the data to keep the estimate of the intrinsic ability of species to produce offspring fixed. The results do not change qualitatively (i.e., when the inferior competitor arrives first, it is predicted to coexist with the superior competitor). Second, in response to a comment by Reviewer 1, we have clarified in the response letter and manuscript why the system we are using is appropriate to test novel predictions from the Modern Coexistence Theory, even though we do not empirically assess populations at equilibrium.

Please find the numbered responses to each comment in blue in the response to the reviewers file.

We hope that you will find the revised version of our manuscript worthy of publication in Ecology Letters.

Sincerely yours,

Inês Fragata (on behalf of all authors)
Centre for Ecology, Evolution and Environmental Changes, Faculty of Sciences, University of Lisbon, Campo Grande 1749-016 Lisboa, Portugal

###########################
Response to the reviewers.

Editor
Editors Comments for the Author(s):
This manuscript uses a clever experimental system to test the operation of priority effects within the quantitative framework of modern coexistence theory. Despite the importance Jim and interest of this effort, two reviewers have raised substantial concerns that would need to be

satisfactorily addressed for this manuscript to become suitable for Ecology Letters. These include (1) justification for allowing lambda (population growth in the absence of competition) and intraspecific competition to vary across order-of-arrival treatments when there is no onvious biological reason for this; (2) using more broadly accepted definition of priority effects and adjusting inferences and conclusions concordantly; and (3) explaining whether or not resident species were at their equilibrium density, as required by MCT, and if not justifying conclusions that can still be drawn.

R0. Thank you for the positive comments on our manuscript. The concerns raised by the reviewers are all constructive points, which have helped to improve the quality of this revised version. Below, we explain here how we have dealt with these three major points.

(1) The biological reason to vary both lambda and intraspecific competition is because individuals have access to different resources depending on their order of arrival. However, following the comments from reviewer 2 we have estimated niche and fitness differences using a fixed lambda (estimated using the data from all treatments) and the results remain qualitatively the same. Please find a detailed explanation in **reply 2.1**.

(2) We have used a broader definition of priority effects and adjusted several parts of the text to it. Please see our **reply 1.3** for more details.

(3) Although equilibrium densities remain at the core theory of modern coexistence theory, using such approximation enters in conflict with our treatment of order of arrival because it takes more than two days for mite populations to reach an equilibrium. We therefore have taken an alternative approach (i.e., density and frequency-dependent gradients) which is in agreement with theory and has been often used in past work. Finally, following the reviewers' comments, we performed an additional experiment to quantify the magnitude of induced defenses by the studied species, which provided an additional layer of evidence to tease apart whether mites were led to priority effects due to niche preemption or niche modification. We provide details of the approach undertaken in **reply 1.1** and we discuss this topic further in the new version of the manuscript.

Please read our point-by-point letter in which we have addressed in detail these tree major concerns and all specific comments made by the reviewers.

Reviewer: 1

Comments for the Authors
This study uses a simple yet interesting experimental system to investigate the effect of arrival order on species coexistence and exclusion. I think the focus on priority effects is valuable as priority effects remain fairly under-studied despite their potential importance to understanding species interactions. In addition, the effort to link empirical data to the modern coexistence theory (MCT) approach is also valuable, given that MCT continues to receive a lot of interest in the community ecology literature, yet has rarely been applied to cases where priority effects occur, with a few recent notable exceptions that are properly cited in the manuscript. Moreover,

I think that the methods used to estimate MCT's niche and fitness differences from experimental data step-by-step using models 6A, B, C are creative.

R: We thank the reviewer for this encouraging comment and thoughtful review!

Despite these strengths that characterize this paper, I believe there are two major difficulties with interpretation of the analysis presented, which I would need to see addressed in a revised manuscript, whether or not it is to be accepted to Ecology Letters or to another journal.

1.1 - First, I think MCT assumes the resident species has reached an equilibrium density at the time of arrival of the invading species. This was not the case in the experiment, so it seems unclear how relevant or helpful the MCT predictions are to explaining the experimental results. I think the paper should discuss this point carefully.

R 1.1: The reviewer raises an important issue. Indeed, measuring whether a species can invade the resident community when the competitor is at equilibrium is an approach to test theoretical predictions from MCT. However, such approach is rarely used in experimental studies testing MCT prediction for logistical reasons (fungi and bacteria are the only cases where this is more easily done because populations reach an equilibrium point in a reasonable amount of time; Grainger et al. 2019). An alternative approach, which is widely used in the literature (Levine and HilleRisLambers 2009, Godoy et al. 2014, Cardinaux et al. 2018, Bimler et al. 2018), is to estimate the strength of these intra and interspecific interactions by creating a combination of density and frequency-dependent gradients (Kraft et al. 2015, Hart et al. 2018). This approach is effective because it covers a wide range of the density- and the frequency-dependent gradients.

Moreover, a critical aspect for our experiment is that we varied the difference in the order of arrival by two days, but it takes much more than two days for the spider mite populations to reach equilibrium densities (Mitchell 1973). Therefore, if we would have employed the approximation of equilibrium densities, we would not be able to apply the order of arrival treatment of two days. This is why, instead, we varied frequency gradients such as to estimate the ratio of intra versus interspecific competition.

This discussion has been included in lines 181-186. We would like to highlight that both approximations test the same theoretical predictions, but use different approaches.

1.2 - Second, the manuscript places a lot of emphasis on teasing apart niche preemption and niche modification as mechanisms causing priority effects. I understand the logic the paper is trying to use with respect to expected competitive outcomes and leaf occupancy patterns, but I don't think it is possible to say from the data presented that niche preemption rather than niche modification was the main factor as conclusively as claimed in the paper. To strengthen this argument, it would have been helpful if data that quantified induced defense (e.g., chemical changes in the leaves) were provided. Otherwise, it seems difficult to eliminate the possibility that niche modification via induced defense, instead of or in addition to spatial niche preemption, affected mite abundances. The paper should discuss this issue as a major caveat, or present a more convincing case explaining why the conclusion can be drawn so definitely, if that is possible.

R 1.2: Indeed, the discussion about teasing apart between niche preemption and niche modification is a very interesting one. In the invasion experiments, we were not able to estimate the induced defenses because leaves were too overexploited by the mites. Instead, we performed an additional experiment in which we estimated the ability of each of the two populations (to suppress/induce defenses in tomato plants). By quantifying specific defense gene expression, we observed that both species suppress plant defenses to a similar extent. This novel result provides fundamental evidence to tease apart niche modification from niche preemption for the following reasons. If niche modification was affecting the probability of the two species to coexist, we would have seen facilitative effects for both species, i.e. an increase in the growth rate of one species when the other species arrived first. However, we observe the opposite, namely, the species that arrives first shows an increase in growth rate at the cost of reducing the growth rate of the second species. Moreover, we observe a change in leaf occupancy towards the most preferred leaves in species that arrive first. Accounting for all experiments, our findings indicate that niche preemption is the most plausible reason for the change in the competitive outcome of both species depending on the order of arrival. This experiment and the consequences thereof, have been included in lines 249-255, and 355-359, and 386-392, respectively. We have also added a new co-author, Mariya Kozak, who performed the experiment.

In addition, I have the following comments:

1.3 - I believe the definition of priority effects used in this paper is unnecessarily narrow and, consequently, the take-home message is confusing and misleading (e.g., "we did not observe that order of arrival cause priority effects." in the abstract). It is true that priority effects as often discussed in the context of MCT specifically refer only to historically contingent competitive exclusion. However, many authors, particularly those of empirical studies, define priority effects more broadly, to refer to any effect of arrival order on the outcome of species interactions, as indicated in the first sentence of the Introduction of this manuscript (line 62). According to this commonly used definition, differences in arrival order did lead to priority effects in the reported experiment. The narrow definition in the MCT context hinders progress with understanding of historical contingency in community assembly, as recently argued by Song et al. 2021 (see reference below). For these reasons, I think the manuscript should be revised accordingly.

R 1.3: We agree that the definition of priority effects used in Modern Coexistence Theory (MCT) does not encompass all the processes that are involved, and can lead to, priority effects. We have adjusted the results and discussion sections to reflect this gap and we clarified that, although we find priority effects, as in Fukami 2015 and other studies, these are not created by positive frequency dependent processes (as defined by Ke and Letten 2018). This can be seen in lines 329-334, 432-436.

1.4 - Not strong enough evidence is presented in the paper to support the paper title, as I

discussed above. I think the title should be revised accordingly. In addition, "Order of arrival" should be changed to "Specific order of arrival" or similar in the title to make it more precise.

R 1.4: Given the additional evidence found in the experiment that we have incorporated in this new version of the manuscript, we are confident that the conclusion stated in the title can be drawn. Therefore, we opted for maintaining the initial title.

1.5 - I believe the results indicate niche differences were reduced (from very large "negative" difference to small, near-zero difference), not increased, when the less competitive species came first, contrary to what is currently discussed in lines 54 and 413. The text should be corrected.

R 1.5: The reviewer is right, and we apologize for the lack of clarity. We have changed the text in lines 55 accordingly.

1.6 - "Facilitative effects are expected to foster coexistence (Fukami 2015)" (line 69) - This is not necessary the case, and I do not think the cited reference indicated that. Some facilitative priority effects can prevent coexistence (for example, when early-arriving species facilitate, and subsequently get excluded by, late-arriving species).

R 1.6: We apologize for the misinterpretation. We have changed the sentence to reflect that such facilitative effects can both promote or limit coexistence between species (cf. lines 70-73).

1.7 - Line 104. I think it is strange to say that spatial distribution is a trait. Behaviors and other species characteristics that lead to a certain spatial distribution can be traits, but not spatial distribution itself.

R 1.7: We have changed this definition (lines 104-105).

1.8 - Line 369 "These multiple lines of evidence challenge current paradigms on the direct connection between order of arrival and priority effects." I think this sentence should be deleted because it is not true. The sentence may be true if we use a very narrow definition of priority effects, but it is not a common definition, as I discussed above.

R 1.8: As mentioned above, we now interpret our results in relation to both the broader definition of priority effects and modern coexistence theory. We changed this sentence to reflect that our results challenge the direct connection between (inhibitory) priority effects and coexistence (lines 372-373).

1.9 - Priority effects have been examined on spider mites at larger spatial and temporal scales in the field (Hougen-Eitzman and Karban 1995). I think this reference should be cited.
R 1.9: We thank the reviewer for pointing out this reference! We have added it in lines 109-112.

Reviewer: 2

Comments for the Authors
In the manuscript, "Order of arrival promotes coexistence via spatial niche preemption by the weak competitor" Ines et al. use a study of the assembly dynamics of two species of <i>Tetranychus</i> mites to explore the relationship between the predictions of modern coexistence theory and species arrival order and frequency. Overall, the paper is very well written and was a pleasure to read. The ideas explored in the paper are also novel and could contribute to the much needed bridge between the theoretical predictions of coexistence theory and the historically contingent dynamics that are common in ecological systems.

R: We thank the reviewer for these positive and thoughtful comments.

2.1 - My one major concern with this study is whether the authors may be misapplying the coexistence model. The authors interpretation of the coexistence model allows the parameters to be dramatically modified by arrival order. Although this phenomenon is interesting, it is not clear to me that the constraints model can actually accommodate this. For example, the analysis that has been presented allows lambda (the intrinsic growth rate in the absence of competition) to vary among treatments. How can growth rate in the absence of competition be modified by arrival order, which implies the presence of species interactions/competition? Similarly, why would arrival order affect the intraspecific competition coefficients? A much more thorough accounting of the implications of allowing these parameters to vary is needed at the very least. However, my intuition is that the coexistence model that is currently used in the manuscript simply can not properly describe the phenomenon being presented without modification. Or possibly explicit reinterpretation of the parameters.

R 2.1: The reviewer made a very important point, and we agree it deserves a detailed justification. Initially, we estimated the models with lambdas that varied depending on the order of arrival because we considered that because individuals arrive at different timings, they will have access to different resources. Nevertheless, we fully agree that without a specific test to show if the intrinsic growth rate changes with the order of arrival, we should assume that the intrinsic growth rate is not affected by it. Hence, we have decided to follow the reviewer comment and estimated a single lambda. To this aim, we combined data from all our experiments and applied the nested modelling approach described in the revised methods section. These changes were added to the Methods section in lines 213-219 and in the supplementary material and methods (lines 87-95).

Although we modified the way we estimated lambda according to the reviewer´s comment, we have not done so for estimating per-capita interaction strengths. We firmly believe that intraspecific competition is modified by the order of arrival because when species arrive first there are more resources and more locations to disperse to, and therefore, the per capita effect on one individual limiting other individuals of the same species is weaker. Conversely, when they arrive later, intraspecific competition is stronger for the simple reason that there are less

resources and less locations to disperse to. Thus, to sum up, we maintained our initial approach to estimate intra and interspecific competitive ability for each treatment but revised the calculation of lambda as suggested by the Reviewer.

With these new analyses, it is worth highlighting that the results remain very similar to what we report in the initial submission. The only change is that in the treatment in which *T. evansi* arrives first, we now see a small overlap of the confidence interval with the priority effects region in the coexistence plot. We added a brief mention of this in lines 324-326.

2.2 - The authors also suggest in the introduction that facilitative dynamics are likely in the system because one species suppresses plant defense. This is also something that the coexistence model can not accommodate.

R 2.2: The reviewer is correct that the definitions of niche and fitness differences and therefore coexistence conditions do not allow for facilitative interactions. We have modified this sentence to avoid potential confusion.

2.3 - None of the analyses appear to account for the fact that there was variation in the order in which the leaf pairs were added to the boxes. Given that in some reps leaves 2-4 were added before 3-5, and sometime the opposite, does it really make sense to view leaf age/position as a continuous/ordinal? The mites are not actually choosing among all leaf ages during the initial assembly. The choice is really for a younger vs older leaf in the environment. And perhaps the priority effect is stronger in the boxes were leaves 2-4 were added first because there is a bigger differences in quality between leaves 2-4 than 3-5? Surely if leaf age matters, then leaf of a given position/age added at the beginning vs after 2-weeks is different.

R 2.3: Indeed, this could potentially bias our results. Because we were aware of this issue, we ensured its mitigation by performing replicates with either leaves 2-4 or leaves 3-5 added first. Thus, if there was an effect due to this variation it would be evenly spread across treatments. Nevertheless, we tested if leaf order affected our results by performing the same analyses separately for the groups of 2-4 vs 3-5 first boxes. Overall, we obtain the same results for each group, although now we see that *T. urticae* clearly prefers leaves that were put in second place, which suggests that they disperse more to leaves that have not been previously colonized. However, as mentioned above, this does not affect the core results observed previously. We added these results as supplementary material, and mention them in the results section, lines 314-315, 346-347

2.4 - Finally, I was not able to assess the specific implementation of any of the analyses that are presented in the manuscript because the git hub repository listed in the methods is not accessible.

R 2.4: Thank you for noticing. It is now open with all the information available.

Other minor comments:

2.5 - Line 49 - "...we show that order of arrival interacts with species' competitive ability to

determine competitive outcomes." I am not sure what 'interacts with' means in this case. I think the authors are arguing that order of arrival change's competitive ability…
R 2.5: Corrected (lines 50-51).

2.6 - 205 - "...negative niche differences…" This is not at all intuitive. I understand how this arises from the model, but a clear description would help many readers.
R 2.6: We thank the reviewer for this comment and accordingly, we added a brief explanation (lines 209-211).

2.7 - Line 228 – leaflet? By line 294 I see the reason for this, but it should be clarified in the data collection section of the methods. Also, a general point: questions about spatial aggregation do not come up until related methods. It would be much easier to follow the methods if a clear statement of each of the specific hypotheses that will be tested was presented in the first paragraph of the methods or at the end of the introduction.
R 2.7: We removed the leaflet information, as the meaningful information is that on the number of individuals per leaf. We added a brief explanation of the different hypotheses tested in the beginning of the experiment section to clarify the methods used.

2.8 - Line 229-223 Why is the intrinsic growth rate being measured under different conditions than the rest of the experiment?
R 2.8: A separate experiment was done because it was not possible to extract values of the intrinsic growth per female from the invasion experiment, as several individuals were introduced simultaneously, and their fecundity could not be tracked.

2.9 - Line 246 – 7 levels? 3 initial frequencies * 3 arrival orders = 9. I see 9 levels in Figure 1, but there is no data for 2 of the level on the x-axis? Maybe these treatments were not actually included? That would explain my confusion about 7 vs 9 treatment levels. But then why are they shown on this figure? This need to be better clarified in the initial description of the experimental design and unused treatments should most definitely not be show in the figure.
R 2.9: Good point. Due to logistical limitations (i.e., lack of space and people), we did not perform the complete orthogonal design. Thus, there is no treatment that involves 19 mites invading one mite established for 48h. We have now removed the unused levels from the figure and added a supplementary figure detailing the experimental setup for this experiment. We hope these modifications serve to better explain the experimental design.

2.10 - Lines 265-269 Am I correct in understanding that the single-species (leaf-disk) experiment was only used to get an initial estimate of lambda that was used to initialize the more complex models which were fit to the multi-species experimental treatments? I think this must be true because the estimates of lambda vary (Fig. S1). However, isn't lambda supposed to represent the intrinsic growth rate of a species in the absence of competition? If so, why would it vary under different arrival order treatments?
R 2.10: We apologize for the lack of clarity. The leaf disks were used to obtain an initial estimate of the lambdas, which was finally assessed in combination with the competition experiment information through a series of nested models. Although we obtained variation

across individuals in offspring production, we used all this variation to estimate an average lambda value and its standard error to later compute niche and fitness differences and the likelihood of species to coexist.

2.11 - Line 468-248 "This allowed performing a priori contrasts to compare between different orders of arrival and frequencies for each species, since our experimental design was not orthogonal." Why does merging the 2 dependent variable into a single factors help here? Why not compare estimated marginal means from the model that reflects the actual experimental design?
R 2.11: To maintain the same logic applied to the rest of the statistical analyses, we now use a single factor (Treatment) and test the impact of frequency and order of arrival by performing contrasts between treatment levels.

2.12 - Do the predicted mechanisms of plant defense suppression or stimulation apply to detached leaves?
R 2.12: Good point. Following the comments of the first reviewer, we have performed an extra experiment to test plant defense induction and suppression. This new experiment shows that there is suppression and induction in detached leaves.

2.13 - Line 324 - "...Tetranychus evansi (the superior competitor) excluded T. urticae when it arrived first or at the same time." But the results don't show that T. urticae was excluded. I am confused by this statement. Do you mean that T. urticae populations are predicted to be eventually excluded?
R 2.13: The reviewer is right, and the text has been changed in line 321.

2.14 - Figure 2 – Are the uncertainties in parameter estimates show in Fig S1 propagated into this figure? The differences among arrival order treatments seem much more extreme in Fig 2 than Fig S1 would suggest. Maybe I am missing something?
R 2.14: We use the upper and lower confidence limits of intrinsic growth rate and intra and interspecific competitive coefficients estimated from the models to calculate the upper and lower limits of niche differences and fitness differences.

2.15 - Figure 3 – Why are no treatments with T. evansi first shown?
R 2.15: All treatments are shown in the supplementary figure S2, we just choose the comparison between *T. urticae* arriving first and at the same time, as this was the most striking observation.

References

Bulleri, F., Bruno, J. F., Silliman, B. R., & Stachowicz, J. J. (2016). Facilitation and the niche: implications for coexistence, range shifts and ecosystem functioning. Functional Ecology, 30(1), 70-78.

Bimler, M. D., Stouffer, D. B., Lai, H. R., & Mayfield, M. M. (2018). Accurate predictions of coexistence in natural systems require the inclusion of facilitative interactions and environmental dependency. Journal of Ecology, 106(5), 1839-1852.

Cardinaux, A., Hart, S. P., & Alexander, J. M. (2018). Do soil biota influence the outcome of novel interactions between plant competitors?. Journal of Ecology, 106(5), 1853-1863.

Fukami, T. (2015). Historical contingency in community assembly: integrating niches, species pools, and priority effects. Annual Review of Ecology, Evolution, and Systematics, 46, 1-23.

Godoy, O., & Levine, J. M. (2014). Phenology effects on invasion success: insights from coupling field experiments to coexistence theory. Ecology, 95(3), 726-736.

Grainger, T. N., Letten, A. D., Gilbert, B., & Fukami, T. (2019). Applying modern coexistence theory to priority effects. Proceedings of the National Academy of Sciences, 116(13), 6205-6210.

Hart, S. P., Freckleton, R. P., & Levine, J. M. (2018). How to quantify competitive ability. Journal of Ecology, 106(5), 1902-1909.

Ke, P. J., & Letten, A. D. (2018). Coexistence theory and the frequency-dependence of priority effects. Nature Ecology & Evolution, 2(11), 1691-1695.

Kraft, N. J., Godoy, O., & Levine, J. M. (2015). Plant functional traits and the multidimensional nature of species coexistence. Proceedings of the National Academy of Sciences, 112(3), 797-802.

Levine, J. M., & HilleRisLambers, J. (2009). The importance of niches for the maintenance of species diversity. Nature, 461(7261), 254-257.

Mitchell, R. (1973). Growth and population dynamics of a spider mite (Tetranychus urticae K., Acarina: Tetranychidae). Ecology, 54(6), 1349-1355.

Song C, Fukami T, Saavedra S. 2021. Untangling the complexity of priority effects in multispecies communities. Ecology Letters 24: 2301-2313.